OUTLINE OF THE PROPOSED TOPIC OF RESEARCH

Candidate

Name of the Candidate	:	Prashant Khanna
ID No.	:	2011PhDENGG005
Place of Research Work and Organization	:	JK Lakshmipat University, Jaipur (JKLU)

Proposed Supervisor

Name	:	Dr B.V. Babu
Qualification	:	Ph.D.
Designation	:	Director, Institute of Engineering and Technology
Organization	:	JKLU, Jaipur

1. Proposed Topic of Research

Cloud Service Brokering : Optimization in Provision of Autonomic Cloud Services

2. Objectives of the Proposed Research

(i) To study the role of a Broker Entity for provision of automatic service in the Cloud Framework.

(ii) To implement a cloud service broker featuring autonomic and automatic negotiation of Service Level Agreements (SLA) for service deployment.

(iii) To propose a uniform and standardized abstraction layer for monitoring of provisioned services by a broker operating in a heterogeneous cloud environment.

(iv) To use Evolutionary Computation principles to negotiate and evaluate available resource policies while providing brokering services in the cloud

(v) To study the effect of SLA failures in the cloud service architecture when viewed from the broker's perspective.

3. Background of the Proposed Research

3.1 Introduction:

Cloud Computing ("Cloud") is a term that describes the present evolutionary development of many existing technologies and approaches to computing that at its most basic, separates application and information resources from the underlying infrastructure and mechanisms used to deliver them. In addition, it also provides an elastic scale and the utility model for allocation of the resources available on a virtual platform. Cloud computing, at its core, is a service model that is promising to enhance collaboration, agility, scale, availability and the hope of cost reduction through optimized and efficient computing for users who intend to do business on a collaborative work model (Dollmann, Houy, Fettke & Loos, 2011).

Out of several definitions that exist in literature on what the cloud is and how it can be summarized, perhaps the most apt one is the one provided by National Institute of Standards and Technology (NIST, 2011) where it defines Cloud Computing as "*Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.*" This service model defines the way entities interact with each other and the way components are associated. This study pertains to the most recent player or entity in the NIST model – the Cloud Broker.

At the hardware level, the Cloud describes the use of a collection of distributed services, applications, information and infrastructure comprised of pools of compute, network, information and storage resources. To qualify as the cloud environment, these components should be provisioned to be *rapidly orchestrated, provided, implemented* and *decommissioned* using an **on-demand utility-like model of allocation and consumption**. This ability of the cloud is what it differentiates it from the erstwhile models like Parallel Computing, Distributed Computing, and off late, Grid Computing model.

Cloud services are often utilized in conjunction with, and enabled by, virtualization technologies to provide dynamic integration, provisioning, orchestration, mobility and scale. Cloud computing is changing the way IT services are provisioned **automatically and dynamically in a federated network environment**. An interesting issue of research that emerges in this domain is the optimal way to aggregate, arbitrate and provide interoperable,

but differing services (both hardware and software), to a set of users. Such a framework would have a deep business and technological impact on the way services are presently being used and provisioned in the IT arena.

Brokering Framework. Any electronic service brokering architecture, in general, must have the ability to support a service delivery infrastructure for integration, delivery and management of composite services in a multi-provider heterogeneous networks environment. It is no different in the cloud service provisioning environment. In the present stage of evolution of the cloud as a repository of services, the provisioning is a current topic of research. The cloud paradigm is currently in a state of transition and multiple players are trying to dominate the service delivery scene. The cloud providers are competing with the cloud brokers to deliver the intended service to the cloud consumer, but this model of service interaction is not bearing the desired results due to multiple barriers of scale and other technical and managerial issues. This research on the subject, supported by the industry case studies available through reports, indicate that the player who is likely to emerge as the principal stake holder in provisioning and arbitraging of services as a truly elastic and dynamic package for the consumer would be the cloud broker (Gartner, 2010). Such service provisioning is already appealing to business entrants who are not yet as big as Google or Amazon, but have the understanding of how the cloud works (Cheng, Gottlieb, Levin, & Lin, 2011). Forrester (Herbert & Erickson, 2011), in their annual report in 2011, also cite brokering services in the cloud to be the next wave of Cloud evolution. However, the present state of cloud implementation is highly proprietary and private, akin to islands of highly autonomous island solutions which do not have any linking ferry services which can carry the inhabitants across. The cloud brokering service available today is thus confined to a miniscule subset of matching services that are seamlessly able to speak to each other. There is a serious void in interoperability between cloud solutions that are not being addressed by the present generation of brokering service providers, either due to technological incompatibilities or due to economic and managerial issues.

Though Cloud computing is a highly studied topic today and a large body of research has gone into studying specific standards of interoperability amongst clouds and how they are to be achieved, the aspects of brokering services to the end client from amongst those available is finding refereed status only recently (Yfoulis & Gounaris, 2010; Jrad, Tao, & Streit, 2012; Khanna & Babu, 2012). As per NIST (NIST, 2011) cloud broker has been described as an

entity that manages the use, performance and delivery of cloud services and negotiates relationships between cloud providers and cloud consumers. The methods and processes which enable a broker to optimally achieve this as an autonomic and seamless service is a matter of research.

Cloud service providers offer management services to manage the cloud service implementation remotely. The Institute for the Future (Gartner, 2010) reported that server farms would grow to support supercomputing applications as more organizations outsourced computing capacity to commercial server farms (PA Consulting Group, 2010). It is perceived by the researcher that the need for such server farms to provide seamless and dynamic service to the intended providers and scale to the desired levels would be possible only through the active involvement of the entity of a Cloud Broker.

The reference architecture in the latest NIST document (NIST, 2011) highlights interactions amongst entities defined and provides a companion cloud computing taxonomy detailing the definitions and relationships of a control vocabulary. The document identifies five major actors to enable the reference model to work, namely, **cloud consumer, cloud provider, cloud carrier, cloud auditor, and cloud broker**. Each actor is an entity (a person or an organization) that participates in a transaction or process or performs tasks in cloud computing. Each of the players have been defined and redefined in literature and the use case(s) to make the model successful has also been commented upon extensively. Amongst the actors defined in the NIST model, the cloud broker was an add-on after much thought. Gartner, in a report in 2011 (Gartner, 2011) indicated that cloud brokering services in the cloud service marketplace is emerging as a promising low-risk business model for offering new and value-added services through cross provider service delivery and partnership. This assertion has made a major impact on the industry as well as the academia. This study utilizes the NIST model as the framework for proposing and building a Semantic Brokering service that provides automated service provisioning and service monitoring. The role of the broker has been discussed in a recent paper by the researcher (Khanna & Babu, 2012).

The study is based on acting and adopting the role of a Broker, in simulation, while ensuring that the cloud provider provisions the required services to the cloud consumer.

3.2 Literature Review of Research Topic

Gartner (Gartner, 2010) anticipates that large enterprises will soon have a dynamic sourcing team responsible for ongoing cloud sourcing decisions and management. It is no longer a question of whether or not users will *use* cloud computing; organizations, both small and large, must ensure that they are cloud-ready and should evaluate opportunities by applying a structured framework to identify potential candidates for cloud-based services. It is clear that cloud computing represents a business model revolution rather than a technology innovation (Accenture Technology Labs, 2010; Deloitte, 2010; Technology Review, 2010).

Virtualization is the underlying platform on which most of the cloud services are based. (Tanaka, Tarui, & Naono, 2009) analyzed that server virtualization is now required for data center systems to reduce the number of servers. Businesses aim to reduce costs while making the most of available resources by mapping their needs and requirements of computation and hardware. The cloud paradigm enables businesses to achieve this in a dynamic and elastic manner. The cloud brokering framework functions in the middle of the cloud provider and the cloud consumer to enable this dynamic provisioning. The position of a cloud broker in the model for provisioning of services and with the special emphasis of provision of automatic and dynamic services in the environment has made the concept of service brokering in the cloud the subject of many researches presently underway in the academia. This entails creation of toolkits and frameworks that can emulate and cloud environment, while saving costs on actually hiring and using services from providers.

A well-known research project is the Cloudbus toolkit (Buyya, Pandey, & Vecchiola, 2012) that defines a complete architecture for market-oriented cloud computing. Buyya et al have defined the three key components of this architecture as the **Cloud Broker, a Market Maker and an InterCloud** (Buyya, Ranjan, & Calheiros, 2010). The Cloud Broker schedules applications on behalf of the user by specifying the desired Quality of Service (QoS) requirements, whereas the Market-Maker acts as a mediator bringing together Cloud providers and customers. It aggregates infrastructure demands from the Cloud Broker and matches them against the available resources published by the Cloud providers. The InterCloud provides a scalable federated computing environment composed of heterogeneous interconnected Clouds enabling the Intercloud resource sharing. This architecture is still under development at MIT and IISc, Bangalore and several instances of the architecture are available at various universities for development, including an internal Laboratory at the researcher's internal cloud premise, where the present research setup will be initiated.

However, experiment with the aforementioned results with Aneka (Aneka enterprise Cloud platform, 2012) and Amazon EC2 (Amazon Elastic Compute cloud EC2, 2012) based Clouds demonstrated that the market-oriented Cloudbus architecture brings benefits to user's application performance in optimizing the cost and execution time.

CloudAnalyst (Wickremasinghe, Calheiros, & Buyya, 2010) is another graphical simulation tool built on top of the CloudSim toolkit (CloudSim Toolkit 2.1.1, 2012), developed by the Cloud Computing and Distributed Systems (CLOUDS) laboratory at University of Melbourne to model and analyze the behavior of large social network applications. The Internet traffic routing between the user bases located in different geographic locations and the datacenters, is controlled in CloudAnalyst by a service broker that decides which datacenter should serve the requests from each user base based on different routing policies. The current version of CloudAnalyst implements three different routing policies, which are network-latency-based routing, response-time-based routing and dynamic-load-based routing. A CloudAnalyst simulation case study of the social network application Facebook proved how load balancing managed by a service broker optimizes the performance and cost of large scale Cloud applications. The research results of this simulation study have been cited several times in the NIST (NIST, 2011) framework for Cloud Adoption by users.

The EU funded OPTIMIS (Ferrer, 2012) project drives the development of a toolkit to optimize the full service lifecycle in the Cloud. It's proposed flexible multi-Cloud architecture includes a service broker that allows decision making after taking into account many aspects like trust, cost and risk. Although the toolkit is still not implemented, the conducted simulation experiments with real workload traces prove the benefits from the use of cost and risk aspects as elasticity policies in the decision making.

The work by Kerterz et al (Kertesz, Kecskemeti, & Brandic, 2011) proposed an SLA-based Service Virtualization (SSV) architecture, which is built on three main components: a Meta-Negotiator responsible for agreement negotiations, a Meta-Broker for selecting the proper execution environment and an Automatic Service Deployer for service virtualization and on-demand deployment.

Most existing works in literature primarily stress on using SLAs to guarantee consumer of cloud services a level of performance, that is defined by abstract metrics, directly from the cloud service providers to the end client or cloud consumers (Alhamad, Dillon, & Chang,

2010; Alhamad, Dillon, & Chang, 2011). There is a void in research on SLA formulation strategies between the cloud service broker and the cloud consumer and between the cloud service broker and the cloud service provider in refereed literature. This research is an attempt to highlight the research void and recommend a framework which can be developed for creation of enforceable and implementable SLAs in the cloud paradigm. One of the objectives of this research emanate from the present void in defining the metrics for adoption of service assurance in the cloud perspective.

The architecture of the cloud, whether public, private, community or hybrid, would make it non trivial to propose and implement binding frameworks in the absence of accurate measuring and monitoring mechanisms for service assurance. This is especially true for a use case when the broker is aggregating and arbitraging services from multiple heterogeneous cloud service providers and packaging them as a service bundle for the end client. Previous work on the subject include (Lamanna, Skene, & Emmerich, 2003; Chhetri, Quo, & Kowalczyk, 2012) and (Chhetri, Quoc, Kowalczyk, & Cam, 2011) that pertains to SLA formulation, but does not address the aspects of the cloud brokering actor's role in the provisioning of services. Alhamad et al in (Alhamad, Dillon, & Chang, 2010) and (Alhamad, Dillon, & Chang, 2011) discuss the aspect of SLA and performance measurement in his recent findings but does not address the issue in the perspective of how a broker would become a party to the SLA agreement between the end user or the cloud consumer and the cloud service provider. Most recent refereed work pertains to the relationship between the cloud provider entity and the cloud consumer with frameworks designed to emulate the relationship. However these frameworks do not illustrate the aspect of arbitraged and aggregated services for a service assurance model from the perspective of a cloud broker. This research is an attempt to address the research void.

A recent work by Wang et al. (Wang, Zuang, Ding, Zhou, Pei, 2011) addresses the aspects of multi-variable SLA based metrics that manages resource scheduling for application provisioning on the cloud. They also recommend a reputation based system for selecting a cloud provider. In (Venticinque, Aversa, Martino, Rak, & Petcu, 2010), Salvatore et al. discuss a framework for broker assigned SLA management service where a high level abstraction model has been recommended. They recommend an architectural design for a system named Cloud Agency that aims to respond to the need for Resources management and offers added value to the existing Cloud services. The proposed system is in charge of

brokering the collection of Cloud resources from different providers that fulfills requirements of user's applications as a *best effort* service. The user is able to delegate to the Agency the necessary checks of the agreement fulfillment, the monitoring of resource utilization and eventually necessary re-negotiations. Balakrishnan and Somasundaram propose a broker framework where SLA enabled broker evaluate the number of resources available in the environment and the number of policies per resource that need to be implemented (Balakrishnan & Somasundaram, 2011). The results presented in the paper indicate that the inclusion of SLA affects the resource selection behavior of the broker. The paper is however silent on the methods to control the affect using an SLA. It does however indicate that the overall performance of the system improves in terms of job throughput with an extra overhead in request processing due to the presence of a broker.

A number of publications, post 2010 (Alhamad, Dillon, & Chang, 2010; Dillon, & Chang, 2011; Wang, Zuang, Ding, Zhou, Pei, 2011; Yfoulis & Gounaris, 2010; Macas, Fit, & Guitart, 2010) are either addressing the aspect of SLA management for brokering services at the level of a resource scheduler, or abstractions of the same when lifted from the grid computing era. The real world applications are viewing SLA performance management and service assurance and provisioning as a combination of availability parameters and associated factors. The carry forward of concepts of web service based SLAs in literature is also evident while drafting cloud based SLAs in recent papers. However, this research on the topic indicates that the business model of provisioning of these two frameworks is very different and mapping the two under the same head would be a mistake. The same has been asserted by various documents published by agencies akin to NIST (NIST, 2011). Quantifiable system level metrics like QoS, CPU utilization, assured storage space, scale up and scale down time in terms of elasticity of service, besides some metrics of security also find mention in industry white papers when they refer to enforceable SLAs. Recent literature also highlight the abstract and non quantifiable aspects of performance management and binding of service issues by cloud service brokers while terming the environment of cloud computing turbulent (Emeakaroha, Brandic, Maurer, & Breskovic, 2011).

3.3 Advantages for an Organization to Adopt the Cloud Computing Service Architecture

There are multiple agencies that are not convinced of the utility of the Cloud Computing Model of service provisioning. While the academia is flush with use cases that state the time for maturing of all technologies that come clubbed under the gamut of "Cloud" have come,

some still doubt the veracity of the industry reports and case studies in public domain as another avatar of technologies that have found flavor recently. The researcher is convinced that the age of the Cloud has arrived. The Business Case for organizations adopting the Cloud Computing Paradigm is very strong and the academia is also realizing the same leading to numerous refereed conferences on the same topic. The use of the Cloud, as explained earlier, provides a number of opportunities to organization, small or big. These have been proven in the market and are now illustrated in the bullets below (Padhye & Mao, 2012):

(i) It dramatically lowers the cost of entry for organizations trying to benefit from compute-intensive business analytics that were hitherto available only to the largest of corporations. Cloud computing also represents a huge opportunity to many third-world countries that have been so far left behind in the IT revolution. There is a special avenue available for colleges and universities aiming to conduct large scale experimentations from within the realms of their laboratories.

(ii) It can provide an almost immediate access to hardware resources, with no upfront capital investments for users, leading to a faster time to market in many businesses. Treating IT as an operational expense (in industry-speak, employing an 'Op-ex' as opposed to a 'Cap-ex' model) also helps in dramatically reducing the upfront costs in corporate computing.

(iii) The cloud becomes an adaptive infrastructure that can be shared by different end users, each of whom might use it in very different ways. The users are separated from each other, and the flexibility of the infrastructure allows for computing loads to be balanced in a dynamic manner as more users join the system. This possible because the process of setting up the infrastructure has become standardized and adding computing capacity has become almost as simple as adding building blocks to an existing grid.

(iv) Cloud computing can lower IT barriers to innovation for organizations.

(v) Cloud computing makes it easier for organizations to scale their services – which are increasingly reliant on accurate information – according to client demand. Since the computing resources are managed through software, organizations can deploy them very fast as new requirements arise. In fact, the goal of cloud computing

is to scale resources up or **down dynamically through software APIs** depending on client load with minimal service provider interaction.

Cloud computing also makes possible **new classes of applications and delivers services** that were not possible before using the Grid, Parallel or Distributed computing. Examples include (Padhye & Mao, 2012):-

(i) Mobile interactive applications that are location, environment, and context-aware and that respond in real time to information provided by human users, nonhuman sensors (e.g. humidity and stress sensors within a shipping container) or even from independent information services (e.g. worldwide weather data)

(ii) Parallel batch processing, that allows users to take advantage of huge amounts of processing power to analyze terabytes of data for relatively small periods of time, while programming abstractions like *Google's MapReduce* (Dean & Ghemawat, 2004) or its open source counterpart *Hadoop* (Shvachko, Kuang, Radia, & Chansler, 2010) makes the complex process of parallel execution of an application over hundreds of servers transparent to programmers.

(iii) Business analytics that can use the vast amount of computer resources to understand customers, buying habits, supply chains and so on from voluminous amounts of data.

(iv) Extensions of compute-intensive desktop applications that can offload the data crunching to the cloud leaving only the rendering of the processed data at the front-end, with the availability of network bandwidth reducing the latency involved.

3.4 SLA in the Cloud

Cloud computing brings in a novel paradigm to foster IT based service economy in scalable computing infrastructures by allowing **guaranteed, on-demand resource allocation** with flexible pricing models. Scalable computing infrastructures not only require autonomous management abilities but also the compliance to users' requirements through SLAs. Such infrastructures should automatically respond to changing components, workload, and environmental conditions as well as prevent violations of agreed SLAs. The essential requirements for SLA-based orchestration of services include agile component-based infrastructure to support these orchestrations; proactive validation of SLAs to prevent

violations; and business enabling requirements including trust, privacy and breach management to address penalty enforcement, renegotiation and recovery.

Flexible and reliable management of SLA agreements is of paramount importance for both Cloud providers and consumers. On the one hand, prevention of SLA violations avoids penalties providers have to pay in case of violations and on the other hand, based on flexible and timely reactions to possible SLA violations, user interaction with the system can be minimized, which enables Cloud computing to take roots as a flexible and reliable form of on-demand computing. An envisaged architecture would mainly focuses on separation of concerns, related to SLAs and services on the one hand, and to the specific domain (e.g., business, software, and infrastructure) on the other. An entity called as a Service Managers would be responsible for all management activities directly related to services. This service responsibilities would includes

- The management of information about available services,
- *The supported types of services,*
- Their offered functionality and their dependencies.

SLA Managers are responsible for all actions that are related to the service-level agreements. They are involved in the negotiation with customers and they are responsible for the planning and optimization of new services that are to be provisioned. Such service managers are the equivalent of Brokers in the present research. Furthermore, they monitor the terms a provider and customer have agreed upon and react in case of violations. The provisioning of a service is a joint effort of all SLA Managers and Service Managers involved. In order to support the metrics decided for a Broker to function, multiple SLA managers and multiple Service Managers should be able to collaborate inside the framework as well as across framework boundaries. The proposed framework for this study would be aiming to provide the following characteristics for the SLA framework :

- **Predictability and Dependability**: The quality characteristics of services can be predicted and enforced at run time.
- **Transparent SLA management**: Service level agreements (SLAs) defining the exact conditions under which services are provided/consumed can be transparently managed across all entities in the cloud stack defined by NIST.

- **Automation**: The whole process of negotiating SLAs and provisioning, delivery and monitoring of services will be automated allowing for highly dynamic and scalable service consumption.

3.5 Evolutionary Computation and SLA Management

In view of the large number of parameters involved and multi options available, with little priori information for solving the issues of selecting resources and hardware to attempt a problem, there is a need to evolve or select a way to address this in real time. Evolutionary algorithms are likely candidates to achieve solutions to such multi-variant optimization problems (Godfrey & Babu, 2004).

We assume that evolutionary computing (EC) is one of the key technologies that can help meet some of the challenges of autonomic computing (Eiben & Smith, 2003; Bäck, Fogel, & Michalewic, 1997). EC is widely applicable, it requires almost no assumptions about the problem to be solved, an evolutionary solver can be usually developed with limited efforts and it produces good quality solutions at acceptable computational costs under a wide range of circumstances. This research will also be an attempt to validate some of the assumptions of EC when used in a multivariate problem which cloud brokers experience while selecting platforms, projects and also solutions to them.

EC encompasses a variety of so-called evolutionary algorithms that all share a common underlying idea: given a population of individuals, the environmental pressure causes natural selection (survival of the fittest), which causes a rise in the fitness of the population over time (Deb, 2005). The main principle behind evolution, be it natural or computer simulated, can be summarized as follows.

If a collection of objects satisfy the following conditions (Gujarathi & Babu, 2009) :

(i) they are able to reproduce

(ii) their offspring inherits their features

(iii) these features can undergo small random, undirected variations

(iv) these features effect their reproduction probabilities

then the features of these objects will change over time in such a way that they will fit their environment better and better. In a formal setting, the environment is represented by a given quality function to be optimized. The population is created by randomly generating a set of candidate solutions, i.e., elements of the function's domain, and the quality function is used as an abstract fitness measure (the higher the better). Based on this fitness, some of the better candidate solutions are chosen to seed the next generation by applying recombination and/or mutation to them.

It is easy to see that ECs fall in the category of generate-and-test algorithms (Godfrey & Babu, 2004). The evaluation (fitness) function represents a heuristic estimation of solution quality, and the search process is driven by the variation and the selection operators. Evolutionary algorithms possess a number of features that can help to position them within in the family of generate-and-test methods (Gujarathi & Babu, 2009):

(i) EAs are population based, i.e., they process a whole collection of candidate solutions simultaneously

(ii) EAs mostly use recombination to mix information of more candidate solutions into a new one

(iii) EAs are stochastic

Recent studies on EC reflect its utility in multiple different domains using Evolutionary techniques like Genetic Algorithms, Swarm Intelligence (Ant Colony Optimization, Particle Swarm Optimization), Differential Evolution, Artificial Life, Learnable Evolution Models etc. Some of the applications that find EC and Evolutionary Techniques useful, especially in multi-objective optimization problems (MOOP), include vehicle routing problems, heat analysis in a chemical plants (Godfrey & Babu, 2004; Babu B.V., 2007; Deb, 2001), designing of gear trains and pressure vehicles, design if shell-and-tube heat exchangers in chemical plants (Babu & Munawar, 2007), design in fluid pumping systems etc. Optimization techniques in Chemical Processing have found several uses in industry in recent times (Babu, Angira, & Rakesh, 2006; Babu & Khan, 2007). A recent development that has found a lot of mention in academia pertains to bio-inspired algorithms being used in optimizing industrial processes. Recent examples of these include (Babu & Gujarathi, 2010).

Babu and co workers, have applied Differential Evolution successfully for both single-objective and multi-objective optimization problems covering a wide variety of complex and non linear applications (Babu & Sastry, 1999; Babu, Chakole, & Mubeen, 2005; Babu, Chakole, & Mubeen, 2005; Babu & Angira, 2006; Angira & Babu, 2006; Angira & Babu, 2006; Babu, Angira, & Rakesh, 2006; Babu, Mubeen, & Chakole, 2007; Babu & Khan, 2007; Babu & Munawar, 2007; Sheth & Babu, 2009; Gujarathi & Babu, 2009; Gujarathi & Babu, 2009; Gujarathi & Babu, 2010; Gujarathi & Babu, 2010; Gujarathi & Babu, 2011; Kumar, Datta, & Babu, 2011).

Brokering frameworks in the Cloud Computing paradigm are highly non linear and multi-variant scenarios and finding global optimal solutions are hard problems to solve. EC based solutions appear likely candidates, especially for multiple objective and threshold mapping problems involving more than two players. This research will attempt exploring the application of Stochastic Optimization techniques for finding solutions to the research void that exists in the area of seamless and optimal Cloud Computing based Brokering Frameworks. It is the researcher's conviction that traditional optimization techniques can at best map the usage to static thresholds, but to determine optimal or near optimal solutions for problems involving assigning resources to the cloud consumer, in near real time is possible only through EC techniques. Recent application cited above in multiple fields bear testimony to this assertion and this research would be a step in exploring this aspect further.

In this research it will be assumed that the brokering service, which the researcher will simulate, is guided by EC principles and the selection process of the resources to be provisioned will be attempted using one of the available EAs. The aim would be to arrive at optimal distribution of resources in near real time to a host of consumers and least cost to the providers through the method of aggregation and arbitration conducted by the Broker.

3.6 Motivation

The motivation for attempting this research is a business case scenario that was experienced while the researcher was managing a cloud based data centre and the requirement was to *emulate a service provider who is enabled to offer services with differentiated, dependable* and adjustable SLAs, and can negotiate concrete SLAs with (individual or groups of) customers in an automated fashion. The business goal imposed additional requirements on *the software providers (to provide components with predictable non-functional behavior) and* infrastructure providers (to support an SLA aware management of resources). The

requirement was also to map the overarching challenge for a service-oriented infrastructure that supports consistent SLA management across all layers of an IT stack and across the various stakeholder perspectives, i.e. the cloud service provider, the cloud service consumer and the cloud service broker. It is also noteworthy to appreciate that the SLA characteristics span across multiple non-functional domains such as security, performance, availability, reliability.

As the use case scenario of the desired system would typically consist of a complex, layered systems, user-level SLAs cannot be directly mapped onto the requirements of physical infrastructure. Cloud based services are composed of other more fundamental services that could be also provided by external parties. Consequently, a stepwise mapping of higher-level SLA requirements onto lower levels and the aggregation of lower-level capabilities to higher levels is required to be implemented for user level SLAs to be mapped to the infrastructure requirements. The vertical information flow between the layers of the IT stack also needs to be carefully reflective of the service interdependencies as well as the originating computation context while creating and implementing Broker based SLA in the cloud paradigm. In addition to SLAs, the vertical information flow also needs to be covered by monitoring, tracking, and accounting data which must support brokering and negotiation processes at each layer. The overall SLA management process is to include different stakeholders, namely customers, service and infrastructure providers, and also various management steps such as service level assessment, new service contracting and monitoring of delivery of services. These issues make creation and enforcement of SLAs in the cloud a non trivial exercise and is worthy of research.

4. Gaps in Existing Research

Cloud Computing is an emerging field with intense interest in both academia as well as in the industry. The emergence of cloud computing is changing this connect between the client and the organization, between the organization and its members and also between the brokers provisioning the services and the consumer-provider team. This is leading to changed ways of doing business and interactions between entities defined in the NIST framework. It is appreciated that one of the key research voids in the cloud computing paradigm is the availability and formulation of clearly defined frameworks for SLA management. The reason is because in the Cloud, SLAs are service level agreements against customer based agreements in the non cloud ways. With the dynamic provision of hardware and almost static

provision of software, difficulty in determining service quality levels where multiple stakeholders are involved, and the need for automatic acquisition and provisioning of resources, creating and monitoring effective SLAs are an open research challenge.

The present research voids require creation of a semantically rich, policy based framework to automate the lifecycle of virtualized services in the cloud using semantic web languages and implementing automated negotiation processes in the cloud. There is also a requirement of defining the lifecycle of a virtualized cloud based service and mapping the same to an enforceable SLA. There is a need to define and research globally acceptable functional specifications for tasks to be automated in creating, negotiating and implementing dynamic SLAs, defining the cost to benefit ratios, the technical policy constraints and specifications, some human agent policies and service assurance policies in the cloud. Study of EC and evolutionary techniques in the field of cloud based brokering is another area that has been an open field of research in recent times. This study will try and address this void as well.

This study is aimed at plugging the gap and identifying the present voids in the framework to implement enforceable SLAs. The cloud brokering framework using evolutionary techniques is another issue that has not been addressed by the academia in sufficient detail. Brokering frameworks has found recent attraction in literature when BP Rimal et. al. (Rimal, Choi, & Lumb, 2009) have conducted a recent survey of cloud computing technologies and found that evolutionary algorithms are good candidates for implementing resource optimization in Brokering frameworks. Pricing models in the cloud have also found recent mention in literature where evolutionary computation techniques (genetic models) have been used to ascertain the best price point for cloud based interactions (Mario & Guitart, 2011). IN comparison to other aspects of cloud computing, it is felt that there still exists a gap in research on this aspect. This study is aimed at addressing the gap.

5. Methodology

Based on the objectives set for the present study and the background work reported in literature with their limitations, the methodology of the present study is proposed as follows:

Phase 1: Literature Survey

A detailed literature survey will be carried out for obtaining the present state of art on the aspect of SLAs with special reference to its avatars in the cloud computing paradigms. The available semantic languages for defining cloud ontology would be researched and gaps

highlighted. The available frameworks for implementing interoperable services in a dynamic manner would be researched and documented. Continuous literature survey will be carried out to collect the information required during various stages of the proposed research. The source of this literature survey will be various available journals like, International Journal of Ubiquitous Computing, IEEE Press, ACM and IARIA Journals, International Journal of Computer Science and Security, International Journal of Engineering, International Journal of Computer Science, SAGE Journals (Online), National Science Journal, NIST publications and other journals subscribed through the library.

Phase 2: Experimental Set-up

In this phase of the research, an experimental setup on an open source platform – Eucalyptus with the open source cloud management platform Openstack will be created. As the present instances of hardware are within the internal laboratories of the researcher, a simulation environment on the cloud tool CloudSim version 3.0 will be used. CloudSim is a free and open source simulator made available through the University of Melbourne.

The research would involve creating multiple service scenarios and the use case model described in the Motivation section of this report will be used to derive the results. It is anticipated that a mix of 4 different services having 3-7 level parameters would be used to gauge the response of the automated resource matching framework. Different use case scenarios available in literature would be used to compare the performance of the proposed framework. This research treats the issue of cloud based brokering and SLA management as a multi-objective optimization problem and would explore the use of evolutionary techniques to arrive at optimal solutions for the issue.

The Brokering problem in the cloud will be considered as a multi-objectives optimization problem and aspect needing optimization would be gauged by the achievement of the following objectives :

(a) Level of Control of the Brokering Framework – Centralized, Hierarchical, Decentralized

(b) Speed of Execution of service implementation

(c) Need for addition resources in terms of hardware or compute cycles of memory or bandwidth

(d) Stability of the platform using a Cloud Brokering Framework

Optimization of the resources available would be aimed at achieving the objectives mentioned above and it is anticipated that this is possible using some techniques of evolutionary computation, while also enforcing SLAs in the framework.

The study would also necessitate establishing a smaller cloud computing laboratory using a hybrid cloud infrastructure (a mix of a Private and a Public Cloud) for cross checking the simulation results. The hardware platform for the experiments will be based on an Infrastructure-as-a-Service (IaaS) deployment model. The study aims to set up a local lab to ensure cloud brokering service for human and technical management of the services in an automated manner.

Phase 3: Mathematical Modeling

A mathematical model will be formulated using evolutionary computation principles and optimization techniques on the basis of results obtained and results envisioned. It is appreciated that the semantic language for comparison of services arbitraged and aggregated by the broker would require mathematical modeling prior and during the implementation in code. The mathematical treatment of the multi-optimization problem is another open problem that will be attempted while undergoing this study.

Phase 4: Analysis of Results

Analysis of experimental results will start with first the experiment and will be continued till end of the research. From the developed experimental data obtained after the initial service integration is done using the broker paradigm, the semantic language created will be tested under different scenarios and further experimental analysis and selection of interoperable and non interoperable service will be done. The comparison of the results obtained while testing the proposed framework will be done against those available in literature. The comparison to real life implementation of these results would be done after availability of economic support. The broker entity and the relationship between the broker and the consumer-provider entities would be done on the basis of literature review and case studies.

Phase 5: Conclusion and Thesis writing

In this phase, all the experimental and theoretical work carried out in different phases will be documented in the form of thesis.

6. Possible Outcomes

In view of the proposed plan of the research, it is anticipated that the following possible outcomes will emerge out of the research :-

(i) A framework for a standardized, semantic language that can be adopted by cloud brokers to make heterogeneous cloud based services interact with each other.

(ii) Exploration of evolutionary techniques in solving multi-objective optimization based problems for ideal resource usage and achievement of agreed upon service levels (through SLAs) will be tested.

(iii) A framework for automatic SLA negotiation and management in the cloud paradigm.

(iv) A methodology to analyze and possibly mitigate the effect of SLA failure in the cloud paradigm.

(v) Creation of a uniform and standardized abstraction layer for monitoring of services provided by a cloud broker.

Plan of Work

The work-plan for the above-mentioned seven activities is shown in figure 2.

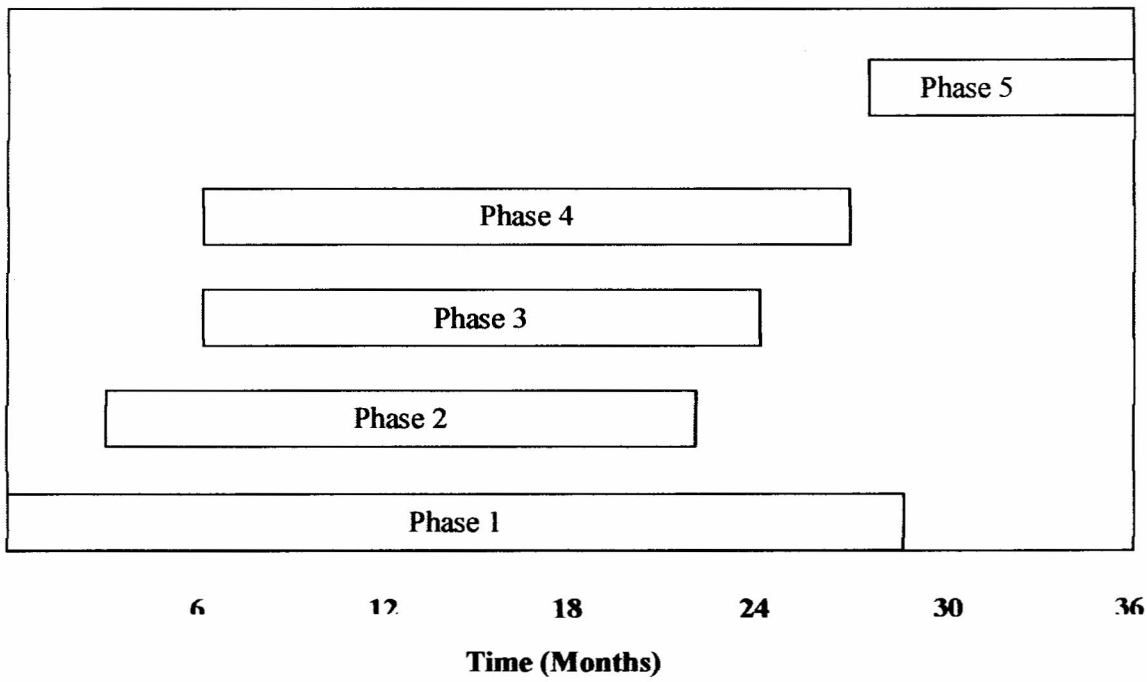

Figure 2. Plan of Work over 36 months

REFERENCES

(2012, Aug 10). Retrieved from Amazon Elastic Compute cloud EC2: http://aws.amazon.com/ec2.

Accenture Technology Labs. (2010). *Everything Elastic - Accenture Technology Vision 2010.*

Alhamad, M., Dillon, T., & Chang, E. (2011). A survey on SLA and performance measurement in cloud computing. *Confederated International Conference on On the Move to Meaningful Internet Systems* (pp. 469-477). Springer-Verlag.

Alhamad, M., Dillon, T., & Chang, E. (2010). SLA-Based Trust Model for Cloud Computin. *13th International Conference on Network-Based Information Systems* (pp. 321-324). IEEE Computer Society.

Aneka enterprise Cloud platform. (2012, Aug 10). Retrieved from Aneka: http://www.manjrasoft.com.

Angira, R., & Babu, B. V. (2006). "Optimization of Process Synthesis and Design Problems: A Modified Differential Evolution Approach. *Chemical Engineering Science , 61* (14), 4707-4721.

Angira, R., & Babu, B. V. (2006). Multi-Objective Optimization using Modified Differential Evolution (MDE). *International Journal of Mathematical Sciences: Special Issue on Recent Trends in Computational Mathematics and Its Applications , 5* (2), 371-387.

Babu, B. V. (2007). Improved Differential Evolution for Single- and Multi-Objective Optimization: MDE, MODE, NSDE, and MNSDE. In B. V. Babu, K. Deb, P. Chakroborty, N. G. Iyengar, & S. Gupta (Eds.), *Advances in Computational Optimization and its Applications* (pp. 24-30). Heidelberg: Springer-Verlag.

Babu, B. V. (2004). *Process Plant Simulation.* Pilani: Oxford University Press.

Babu, B. V., & .Munawar, S. A. (2007). Differential Evolution Strategies for Optimal Design of Shell-and-Tube Heat Exchangers. *Chemical Engineering Science , 62* (14), 3720-3739.

Babu, B. V., & Angira, R. (2006). Modified Differential Evolution (MDE) for Optimization of Non-Linear Chemical Processes. *Computers & Chemical Engineering , 30* (6-7), 989-1002.

Babu, B. V., & Angira, R. (2008). Optimization of Industrial Processes using Improved and Modified Differential Evolution. In *Soft Computing Applications in Industry : Studies in Fuzziness and Soft Computing* (pp. 1-22). Springer Berlin.

Babu, B. V., & Gujarathi, A. M. (2010). Hybrid Multi-objective Differential Evolution (H-MODE) for optimization of Polyethylene Terephthalate (PET) reactor. *International Journal of Bio-inspired Computation.* , *2* (3/4), 213-221.

Babu, B. V., & Khan, M. (2007). Optimization of Reactive Distillation Processes Using Differential Evolution Strategies. *Asia-Pacific Journal of Chemical Engineering* , *2* (4), 322-335.

Babu, B. V., & Sastry, K. K. (1999). Estimation of Heat Transfer Parameters in a Trickle Bed Reactor using Differential Evolution and Orthogonal Collocation. *Computers and Chemical Engineering* , *23* (3), 327-339.

Babu, B. V., Angira, R., & Rakesh. (2006). Performance of Modified Differential Evolution for Optimal Design of Complex and Non-Linear Chemical Processes. *Journal of Experimental & Theoretical Artificial Intelligence* , *18* (4), 501-512.

Babu, B. V., Chakole, P. G., & Mubeen, J. H. (2005). Differential Evolution Strategy for Optimal Design of Gas Transmission Network. *Journal of Multidisciplinary Modeling in Materials and Structures* , *1* (4), 315-328.

Babu, B. V., Chakole, P. G., & Mubeen, J. H. (2005). Multiobjective Differential Evolution (MODE) for Optimization of Adiabatic Styrene Reactor. *Chemical Engineering Science* , *60* (17), 4822-4837.

Babu, B. V., Mubeen, J. H., & Chakole, P. G. (2007). Simulation and Optimization of Wiped-Film Poly-Ethylene Terephthalate (PET) Reactor using Multiobjective Differential Evolution (MODE). *Materials and Manufacturing Processes: Special Issue on Genetic Algorithms in Materials* , *22* (5), 541-552.

Bäck, T., Fogel, D., & Michalewic, Z. (1997). *Handbook of Evolutionary Computation* . Oxford Univ Press .

Balakrishnan, P., & Somasundaram, T. S. (2011). "SLA enabled CARE resource broker . *Future Gener. Comput. Syst*, *23*, pp. 265-279.

Buyya, R., Pandey, S., & Vecchiola, C. (2012). Market-Oriented Cloud Computing and the Cloudbus Toolkit. *Large Scale Network-centric Computing Systems.*

Buyya, R., Ranjan, R., & Calheiros, R. N. (2010). InterCloud- Utility-Oriented Federation of Cloud Computing Environments for Scaling of Application Services. *10th International Conference on Algorithms and Architectures for Parallel Processing* (pp. 13-31). IEEE Press.

Bykov, S., Gelle, A., Kliot, G., Larus, J., Pandya, R., & Thelin, J. (2011). Orleans: Cloud computing for everyone. *In Proceedings of the ACM Symposium on Cloud Computing, Socc'11.*

Cheng, K. E., Gottlieb, Y. M., Levin, G. M., & Lin, F. J. (2011). Service brokering and mediation: Enabling next generation market and customer driven service delivery. *Proc. Tenth International Symposium on Autonomous Decentralized Systems (ISADS '11)*, (pp. 525-530).

Chhetri, M. B., Quo, c. B., & Kowalczyk, R. (2012). Policy-Based Automation of SLA Establishment for Cloud Computing Services. *12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing* (pp. 164-171). Washington: IEEE Computer Society.

Chhetri, M. B., Quoc, B. V., Kowalczyk, R., & Cam, L. D. (2011). Cloud broker: helping you buy better. *12th International Conference on Web Information System Engineering*, (pp. 13-14). Sydney .

CloudSim Toolkit 2.1.1. (2012, Aug 10). Retrieved from CLOUDS Lab, Unversity of Melbourne: http://www.cloudbus.org/cloudsim/

Dean, J., & Ghemawat, S. (2004). MapReduce: simplified data processing on large clusters. *Proceedings of the 6th conference on Symposium on Opearting Systems Design & Implementation* (pp. 10-10). Berkeley: USENIX Association.

Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms* . Kanpur: Wiley.

Deb, K. (2005). *Optimization for Engineering Design.* Kanpur: PHI, Pvt Ltd.

Deelman, E., Singh, G., & Livn, M. (2008). The cost of doing science on the cloud: The montage example. *Proceedings of the 2008 ACM/IEEE Conference on Supercomputing*, (p. 12).

Deloitte. (2010). *Depth perception: A dozen technology trends shaping business and IT in 2010* .

Dollmann, T., Houy, C., Fettke, P., & Loos, P. (2011). Collaborative Business Process Modeling with CoMoMod - A Toolkit for Model Integration in Distributed Cooperation Environments. *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2011 20th IEEE International Workshops*, (pp. 217 - 222). Paris.

Eiben, A., & Smith, J. (2003). *Introduction to Evolutionary Computing* (1st edition, 2003, ISBN: 3-540-40184-9 ed.). Springer.

Eisenhauer, M. P. (2005). Law Issues in Off-shore Outsourcing Transactions. In *Privacy and Security Law Issues.*

Emeakaroha, V. C., Brandic, I., Maurer, M., & Breskovic, I. (2011). SLA-Aware Application Deployment and Resource Allocation in Clouds . *Computer Software and Applications Conference Workshops* (pp. 298-303). IEEE Press.

Ferreiro, D. S. (2010). Guidance on Managing Records in Cloud Computing Environments. *NARA Bulletin* .

Ferrer, A. J. (2012). OPTIMIS: A Holistic Approach to Cloud Service Provisioning. *Future Generation Computer Systems* , 66-77.

Gartner. (2010). *Gartner's Top IT Predictions for 2011-2015.*

Gartner. (2011). *The Future of Cloud Computing.* Gartner.

Godfrey, C., & Babu, B. V. (2004). *New Optimization Techniques in Engineering.* Heidelberg: Springer - Verilag.

Gujarathi, A. M., & Babu, B. V. (2010). Hybrid Multi-objective Differential Evolution (H-MODE) for optimization of Polyethylene Terephthalate (PET) reactor. *International Journal of Bio-inspired Computation* , *2* (3/4), 213-221.

Gujarathi, A. M., & Babu, B. V. (2009). Improved Multi-Objective Differential Evolution (MODE) Approach for Purified Terephthalic Acid (PTA) Oxidation Process. *Materials and Manufacturing Processes* , *24* (3), 303-319.

Gujarathi, A. M., & Babu, B. V. (2011). Multi-objective Optimization of Industrial Processes using Elitist Multi-objective Differential Evolution (Elitist- MODE). *Materials and Manufacturing Processes* , *26* (3), 455-463.

Gujarathi, A. M., & Babu, B. V. (2010). Multi-objective Optimization of Industrial Styrene Reactor: Adiabatic and Pseudo-isothermal Operation. *Chemical Engineering Science* , *65* (6), 2009-2026.

Gujarathi, A. M., & Babu, B. V. (2009). Optimization of Adiabatic Styrene Reactor: A Hybrid Multi-Objective Differential Evolution (H–MODE) Approach. *Industrial & Engineering Chemistry Research* , *48* (24), 11115-11132.

Gunderloy, M. (2008, January 13). Who Protects Your Cloud Data. *Web Worker Daily* .

Heiser, J. (2009). What You Need to Know About Cloud Computing Security and Compliance.

Herbert, & Erickson. (2011). The ROI of cloud apps. *A Total Economic Impact™ Analysis Uncovers Long-Term Value In Cloud Apps* .

Jrad, F., Tao, T., & Streit, A. (2012). SLA Based Service Brokering in Intercloud Environments. *2nd International Conference on Cloud Computing and Services Science* (pp. 76-81). IEEE.

Kertesz, A., Kecskemeti, G., & Brandic, I. (2011). Autonomic SLA-aware Service Virtualization for Distributed Systems. *19th Euromicro International Conference on Parallel, Distributed and Network-based Processin*, (pp. 503-510).

Khanna, P., & Babu, B. (2012). Cloud Computing Brokering Service: A Trust Framework . *he Third International Conference on Cloud Computing, GRIDs, and Virtualization* (p. 7). Nice: IARIA.

Kumar, S., Datta, D., & Babu, B. (2011). Differential Evolution Approach for Reactive Extraction of Propionic Acid using Tri-n-Butyl Phosphate (TBP) in Kerosene and 1-Decanol. *Material and Manufacturing Processes , 26* (9), 1222-1228.

Lamanna, D. D., Skene, J., & Emmerich, W. (2003). Slang: A language for defining service level agreements. *Ninth IEEE workshop of Future Trends of Computing* (pp. 100-106). IEEE.

Macas, M., Fit, J. O., & Guitart, J. (2010). Rule-based SLA management for revenue maximisation in Cloud Computing Markets . *International Conference on Network and Service Management* (pp. 354-357). IEEE Press.

Mario, & Guitart, J. (2011). A genetic model for pricing in cloud computing markets. *Proceedings of the 2011 ACM Symposium on Applied Computing* (pp. 113-118). New York: ACM.

Meiko Jensen, J. S. (2009). On Technical Security Issues in Cloud Computing. *IEEE International Conference on Cloud Computing.*

NIST. (2011). *500-293, US Government Cloud Computing Technology Roadmap, Release 1.0 (Draft), Volume II Useful Information for Cloud Adopters, November 2011.* NIST. Washington: NIST Special Publication.

NIST. (2011). *The NIST Definition of Cloud Computing.* National Institute of Standards and Technology.

PA Consulting Group. (2010). *ICT Horizons: Technology trends for a smarter world, www.paconsulting.com.*

Padhye, J., & Mao, Z. M. (2012). Conference Proceedings . *Proceedings of the third ACM workshop on Mobile cloud computing and services.* Lake District: ACM Press.

Peter Mell, T. G. (2009). The NIST Definition of Cloud Computing, Version 15,.

Rimal, B. P., Choi, E., & Lumb, I. (2009). A Taxonomy and Survey of Cloud Computing Systems. *Proceedings of the 2009 Fifth International Joint Conference on INC, IMS and IDC* , 44-51.

Sheth, P., & Babu, B. (2009). Differential Evolution Approach for obtaining Kinetic Parameters in Non isothermal Pyrolysis of Biomass. *Materials and Manufacturing Processes , 24* (1), 47-52.

Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010). The Hadoop Distributed File System. *Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies* (pp. 1-10). Washington: IEEE Computer Society.

Stokes, J. (2010). EMC's Atmos Shutdown Shows Why Cloud Lock-in Is Still Scary. *Ars Technica* .

Technology Review. (2010). *The Office of the Future,*.

Venticinque, S., Aversa, R., Martino, B., Rak, M., & Petcu, D. (2010). A cloud agency for SLA negotiation and management. *Conference on parallel processing* (pp. 587-594). Springer-Verlag.

Wald, H. (2010). Cloud Computing for the Federal Community. *IAnewsletter* .

Wang, M., X, W., Zhang, W., Ding, F., Zhou, J., & Pei, G. (2011). A Conceptual Platform of SLA in Cloud Computing. *IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing* , (pp. 1131-1135).

Wickremasinghe, B., Calheiros, R. N., & Buyya. (2010). CloudAnalyst: A CloudSim-based Visual Modeller for Analysing Cloud Computing Environments and Applications. *24th IEEE International Conference on Advanced Information Networking and Applications* (pp. 446-452). IEEE Press.

Yfoulis, C. A., & Gounaris, A. (2010). Honoring SLAs on cloud computing services: a control perspective. *Proc. 2nd Workshop on Bio-inspired Algorithms For Distributed Systems* (pp. 29-38). IEEE.